HOW LEARNERS AND FACULTY RATE ONLINE ASSESSMENT METHODS

Fátima Silva, Pedro Reis

Universidade Fernando Pessoa (PORTUGAL)

Abstract

The increasing importance of online education has resulted in attempts to make these teaching/learning means more reliable, valid and feasible in order to avoid the criticisms that are often directed at any e-learning course, above all its assessment process within teaching and learning contexts.

Given the potential of online education to higher education, our question is whether, and in what ways, different types of assessments might have an impact on the view faculty and learners have of the assessment methods used in the Virtual University of Fernando Pessoa University (UFP-UV) in Porto, Portugal, in relation to their reliability, feasibility and validity.

In this paper we will revisit some of the approaches involved in any assessment process, with particular emphasis on online assessment. The case study carried out sought to investigate the extent to which faculty and learners consider the assessment used in the curricular unit of the e-learning course to be feasible, valid and reliable. These impressions were gauged in two different questionnaires (one for the teachers and another for the learners) in which respondents assigned graded quantitative responses to certain statements, and, in the observation section of the questionnaire, they could provide qualitative comments on the question, thus allowing both a quantitative and a qualitative analysis of the data.

The results demonstrate that the virtual classroom is, for both groups in this case study, the mechanism that shows the most reliability whether the virtual classroom acts as a 'real' classroom in which both parties are 'face-to-face' via audiovisual means or only the audio means of the virtual classroom is used to question the learner orally.

Keywords: Online assessment, e-learning, reliability.

1 INTRODUCTION

As the 21st century has heralded a technology-driven society, in the personal, family, educational and professional sphere, education has had to adapt to this new phenomenon in which people, especially the younger generation, are digital natives. Increasingly educational institutions, especially in higher education, are choosing a range of educational processes: they can either continue providing only face-to-face (F2F) courses by continuing with the traditional pedagogic model, or opt for the use of new technologies in blended or full online/e-learning courses.

The increasing importance of online education has resulted in attempts to make this teaching/learning means more reliable, valid and feasible in order to avoid the criticisms that are often directed at any e-learning course, above all its assessment process. If learners perform the activities proposed for the curricular unit, they should be able to demonstrate − in the assessment process − that they have achieved the learning objectives. This process should include the use of a range of instruments as well as ensuring that the learners are engaged in the assessment process which makes them active participants in their learning, thereby making the assessment more reliable and valid. This, in turn, may lead to greater acceptance of e-learning courses by stakeholders and society in general.

2 THEORETICAL BACKGROUND

Assessment is part of our lives as we are continuously weighing/assessing the advantages and disadvantages of decisions we have to make: for instance, what to wear or eat, in the long-term we assess schools to attend, houses/cars to purchase, people who will be our friends, partner, and so on. Fernandes [1] notes that, more or less explicitly and/or more or less formally, assessment is present in all academic domains and in all areas of human activity to ensure that one has quality goods and services which do not pose a risk to one's legitimate interests, health or security. These include

decision making, improvement of procedures and practices, understanding social problems and thus contributing to their possible solutions, understanding the experiences of those involved in social activities, and giving credit to and validating programmes [1].

Assessing is thus a complex social process which involves not only people from different socio-economic contexts, with their own values and practices but also what is being assessed with its own objectives and logic [1]. As assessment is a process developed by people for people, it involves moral and ethical values, personal interpretation, as well as socio-cognitive, socio-cultural and psychological aspects. Thus, the wide range of factors and situations which come into play in the evaluation process make it difficult to expect a consensual acceptance of the result of the evaluation by everyone to whom, in some way, it may be of interest. This point is similar to Boud's [2] idea that assessment is personal and has an impact on everyone, influencing one's learning and 'it helps construct the society of which we are a part. All acts of assessment involve more than is apparent and we must judge them according' (ibid), considering both the immediate short-term needs of certification and the long-term or life-long contribution it can make to the learning process. Furthermore Scalise & Gifford [3] postulate that assessment is 'any interaction with a respondent from which data is collected with the intent of making an inference about the respondent' and it is necessary when learners need accreditation for their learning process. It also measures the outcome of the different participants involved in the learning process, namely, the learner, the teacher/instructor, the course, and the educational institution.

To ensure that the stakeholders accept the curricular unit, the course, or the degree, assessment must be reliable and fair, and beneficial to learning insofar as it has an impact on future [4]. Reliability reflects the degree to which the measurement is accurate and consistent, and can be reproduced in other situations [4]. For assessment to be reliable and fair, different assessors should agree on the mark given to a particular piece of work [5]. Thus, the reliability of an assessment may be affected by several factors such as the assessor's judgments, learner's nervousness, and assessment conditions [6]. Validity shows whether the assessment measures the competences it claims to measure [4]. In relation to this point, Race [5] posits that assessors often do not measure evidence of achievement of intended learning outcomes but rather, for instance, in essays, essay-writing skills, as well as what learners remember about the subject matter, instead of mastery of the subject matter concerned, namely what learners can do with what they processed from the acquired knowledge. In addition to the above criteria for ensuring acceptance of the assessment, Race [5] suggests two others: authenticity, or the real-world relevance of the assessment, refers to 'how well the assessment correlates to the sorts of things students need to be able to do in their career after leaving the educational institution' [5]; and inclusiveness, which answers questions such as, Can this assessment be taken by learners with additional learning needs, for instance, including dyslexia? Does it reduce unfair discrimination?

Thus, the reliability and validity of assessment drive and stimulate learning, providing information to institutions and instructors on the efficiency of the learning process [7]. The role of the learners and the assessors is different: while the former need to know what is expected of them and require feedback in order to improve, those who assess need to ensure that progress is made by the learners, in addition to applying the curriculum, and addressing the requirements of society and stakeholders such as accrediting bodies.

Race [8] states that as the standards involved in assessing learners' work are known by the parties involved, it is necessary to:

> improve assessment to make it more valid, more fair, more transparent to students, better linked to the world outside higher education, and more inclusive so that assessment does not disadvantage students with identified special needs, and enables them to demonstrate their learning in ways where they can show their optimum achievement (p. 73).

2.1 Assessing through e-learning

In the digital environment of the 21st century, in which technology enhances teaching/learning, assessment must reflect the teaching that occurs in the course. In addition, the instructor should use 'multiple measures and authentic assessments' [9] in order to assess if the learners are equipped with the skills required to successfully function in society. This means that:

> there is more likelihood of alignment with outcomes and competencies, a lower possibility that cheating will occur, [...] and an increased likelihood that a true measurement of student

> competency and performance will occur: the use of multiple measures of assessment is simply good pedagogy ([9], p. 27).

According to Eyal [10], learners need to be able to: locate and acquire knowledge independently; wisely use the acquired knowledge to solve problems; make informed choices and critical evaluations; and develop communication and collaboration skills. These should make them 'capable of adaptation and autonomous thinking; in other words − a capacity for self-directed learning that will persist throughout life' ([10], p.42). Authentic assessment in alignment with the teaching process provides evidence of learners' thinking and mastery of the curriculum, and, as a consequence, guides learning. It should focus on cognitive processes which are important in situations outside the particular subject and 'in life outside school [thereby] emulating real life task situations' [11].

Many of the principles applied to assessment concern both F2F and online environments insofar as any good assessment should be learner-centred and not consist solely of tests and quizzes, as was the case in the early days of e-learning. However, these forms of summative assessments were, and still continue to be, 'problematic in that concerns quickly arose regarding the heightened potential for cheating, and snapshots at the end of a module rarely gave insight into the learning' [12].

Despite the existence of LMSs or VLEs in many higher education institutions, Rienties et al. [13] refer that these are often used as a 'simple repository for providing students with access to materials, such as PowerPoint files and reading lists' and tend to be instructor-designated. Thus, to make education more learner-centred, more personalised, flexible and user-centred Web 2.0 technologies are being used. These recognise learners' experiences outside the 'classroom' and communication through social software, such as Facebook or Twitter, and provide them with tools to support their self-governed, problem-based and collaborative activities [14], making the learning process learned-focussed and empowering students to reflect, construct and collaborate while working independently. However, as PLEs are not hosted by educational institutions, the content and maintenance are outside their control, which raises issues of security and reliability at a time when society is still leery of these issues in relation to e-learning

Since assessment is a crucial process of the education process, in addition to being a delicate subject, it is often the part of this process which is most criticised. This is especially true of e-learning due to the absence of F2F interactions that allow instructors to observe the learner, and determine learner response and reaction. Unlike F2F assessment in which the learner is normally 'treated as an isolated individual with limited access to resources and other people' [15], in the e-learning environment, learners have 'access to almost unlimited digital resources' (ibid). As such, assessment methods in an e-learning environment must use multiple measures of learner performance which should be clear, congruent with the learning objectives, in alignment with the subject matter which has been taught and the kind of teaching, and authentic. Assessment must be an ongoing process with both the instructor and the learners being committed to the learning outcomes. It is of utmost importance that the assessment tools and methods safeguard the feasibility and fairness of the teaching/learning process, and ensure appropriate means of administering the assessment. Furthermore, if different assessment tools and activities are used throughout the assessment process, the instructor may be able to have more control over the whole assessment process.

According to Palloff & Pratt [9], a well-designed online course should be learner focused and centred, in which learners are given credit for self-reflection, and it should be incorporated into the design and expectations for the online course. The reflective component of the course should be part of each collaborative activity as learners should be able to reflect on, and then explain, their participation in the activity and their contributions to their peers' work.

Thus, one of the ways to overcome the limitations of online assessment is for the learners to be assessed continuously and formatively using a range of assessment tools, such as self-assessment (fundamental in more constructive approaches); collaborative assessment or peer-assessment; rubrics; quantitative and qualitative participation using synchronous and asynchronous communication and collaboration tools available in online platforms, for instance, chat rooms, e-mail, discussion forums, wikis; quizzes and tests: or even the elaboration of an e-portfolio.

3 METHODOLOGY

Fernando Pessoa University (UFP), a private university in Porto, Portugal, has always aimed to follow its motto 'nova et nove' – teach what is new in a new way. This is reflected in several spheres, for instance, it was the first higher education institution to include the teaching of foreign languages in all

the courses, learners had to use laptops when most people still did not possess this technology, and in 2004, the introduction of Sakai, an open-source Collaboration and Learning Environment (CLE) software, and the creation of the e-learning platform, called UFPUV (Virtual University of Fernando Pessoa University). The UFPUV is popular among instructors and learners as it is used in F2F courses as 'a means to provide learners with extra resources or announcements, as synchronous communication, mainly in language curricular units, and even for testing, through the Test Center' [16].

In order to provide full e-learning courses 'characterised by the use of appropriate educational technologies and by the fact that they have been structured in learning modules, where interactivity, collaborative work and the development of formative activities are essential for the participants' learning' [17], a range of educational technologies are used, mainly Sakai, the UFPUV LMS; Colibri, a Web Collaborative Environment; Educ@st, a service which permits video content; as well as the production of interactive digital contents using Xerte, eXe Learning or CourseLab. At UFPUV, high standards imply that assessment practices comply 'with the learning objectives set out for each curricular unit and give preference to student contributions in acquiring competences' [18]. Means of assessment include diagnostic assessment, process assessments (including formative and summative processes), assessment of synchronous and asynchronous collaborative activities, and final evaluations.

These e-learning courses were provided from 2012 to 2015. Despite their success, they had to be discontinued as a result of Portuguese legislation which restricts fully online courses [19]. Thus, UFP now offers blended courses in addition to the traditional pedagogic courses.

3.1 Aim of the study

Given the potential of online education to higher education, our question is whether, and in what ways, different types of assessments might improve the view faculty and learners have on the assessment methods used in the virtual university of Fernando Pessoa University (UFPUV).

This case study sought to investigate the extent to which faculty and learners consider the assessment used in the curricular unit of the e-learning course to be feasible, valid and reliable. These impressions were gauged in two different questionnaires (one for the teachers and another for the learners) in which respondents assigned graded qualitative responses to certain statements, and, in the observation section of the questionnaire, they could provide qualitative comments on the question, thus allowing both a quantitative and a qualitative analysis of the data.

3.2 Participants and procedure

The sample consisted of 60 instructors of the 67 who teach online courses, that is 89.5% responded to the questionnaires; however, the number of responses from the 185 learners enrolled in UFPUV was low, namely 70,representing 37.8%.

An extensive review of the literature did not render any questionnaire that approached the reliability and feasibility of assessment in this online modality. Thus, the questionnaires were developed based on the literature and the author's own experience of teaching online courses, and validated by conducting focus groups and pre-tests, after which some changes were made [16]. Consent was obtained from the Ethics Commission of Fernando Pessoa University to carry out the questionnaires among the e-learning students and teachers.

Instructors and learners were sent an email explaining the aims of the questionnaires and providing the link to Google Forms. This means of collecting data in Google Drive permits the creation of questionnaires/surveys, which are then automatically recorded and data is managed. In an attempt to have a greater response rate, learners were sent follow-up emails.

4 RESULTS

The questionnaires consisted of two parts: independent variables (including personal and course factors) and dependent variables (attitude toward assessment process and means used in UFPUV). The questionnaire targeting the instructors who teach Curricular Units as e-learning courses aimed to find out, among other factors, the control mechanisms used to ensure that the assessment is reliable and how they classify each of these mechanisms in terms of reliability, and requested instructors to make general comments on making the online courses more reliable.

The questionnaire aimed at the learners studying Curricular Units online at UFPUV had both objective and subjective items. Here we will only refer to one of the items, namely, the control mechanisms that must be used to ensure the reliability of the assessment.

4.1 Instructors' Questionnaires

In their questionnaires instructors were asked about the reliability of the five control mechanisms (Table 1), namely

1 Virtual classrooms with audiovisual means which permit the instructor and the learner(s) to see and hear each other. Thus, the instructor has some control over what is happening during the assessment process.

2 The learners provide oral answers via the virtual classroom; thus, the assessment is as close as possible to a F2F situation.

3 If a time limit should be allocated to each question. This may reduce the possibility of cheating as the learners have a set time to answer the question and may not be able to search the answer online or use other means of getting help.

4 Releasing one question at a time. Each question must be answered online, for instance, in the Assignments area of the Sakai platform within the allocated time. Once again, this may limit the possibility of getting 'help'.

5 Having pre-prepared alternative questions in case of the system fails. If there is a technical breakdown/failure between the instructor and the learner during the assessment, a way of ensuring that there was no cheating in the time span that elapsed may be for the instructor to provide new questions.

Table 1. *The control mechanisms mentioned can ensure reliability of assessment. How do you rate the reliability of each one?*

	Virtual classroom with audiovisual means		**Oral answers via the virtual classroom**		**Time limit for answering a question**		**Releasing one question at a time, instead of the whole test paper**		**Pre-prepared alternative questions in case of system failure**	
Not at all reliable	1	1.9%	0	0%	3	5.5%	4	7.4%	5	9.4%
Not very reliable	18	33.3%	4	7.1%	12	21.8%	12	22.2%	12	22.6%
Reliable	26	48.1%	18	32.1%	35	63.6%	35	64.8%	30	56.6%
Very reliable	9	16.7%	34	60.8%	5	9.1%	3	5.6%	6	11.3%

The 1st column of each point refers to the number of responses while the 2nd is the percentage.

Overall, an analysis of the data indicates that a vast majority of the instructors (over 90%) consider the most reliable control mechanism during assessment to be when the learner replies orally to the question. This may occur due to the fact that the learner is 'face-to-face' − or rather 'camera-to-camera' − with the instructor having to give immediate answers. Even in F2F situations, many education institutions require learners to present oral work in order to show their competence by communicating with the instructor. This almost eliminates the likelihood of cheating and as such is the most reliable for instructors.

Totalling the factors 'reliable' and 'very reliable' shows that in each control mechanism over 60% of instructors rate each as a good means of controlling learners during their assessment. Although it is difficult, in any situation, whether it be in traditional F2F classrooms or online, to prevent cheating during the assessment, the control mechanisms used in e-learning courses of UFP tend to be reliable.

One aspect that stands out from the table is in the use of the virtual classroom: 48.1% consider it reliable but a third of the instructors do not rate it as very reliable. The virtual classroom – despite the limitations because the instructor only has a reduced view of the learner on the other side of the video – allows a more 'personal' and 'real-life' experience of teaching/learning. Assessment through this audiovisual means allows greater control as long as the learner stays online and within the range of

the camera. As such it often makes the instructors feel that they can control the assessment of their learners. However, in this situation a reasonable number of instructors do not believe that they are in control even though they are able to see the students. It would be useful in a future study to find out why so many instructors do not rate the virtual classroom as reliable.

In relation to the qualitative comments, over 50% of the respondents mentioned that the best way to ensure the reliability and feasibility of their assessment is through oral tests.

4.2 Learners' Questionnaires

In relation to the learners' questionnaires, we consider the item which requires respondents to give their opinion on how reliable four control mechanisms are to ensure the feasibility of the assessment process (Table 2). These control mechanisms are the same as those mentioned above for instructors, except for the fifth item that does not appear in their questionnaire.

Table 2. *Which control mechanisms should be used to ensure the feasibility of the assessment process?*

	Virtual classroom with audiovisual means		**Oral answers through the virtual classroom**		**Time limit to answer a question**		**Availability of one question at a time, instead of the whole test**	
Unreliable	0	0%	2	3%	5	8%	9	14%
Slightly reliable	6	9%	6	9%	22	34%	23	35%
Reliable	31	45%	31	47%	27	42%	21	32%
Very reliable	32	46%	27	41%	11	17%	13	20%
1) No. of responses 2) %	1)	2)	1)	2)	1)	2)	1)	2)

The 1st column of each point refers to the number of responses while the 2nd is the percentage.

For the learners, the mechanism that shows the most reliability – around 90% – is the virtual classroom, whether as a 'classroom' with the instructor watching them during the assessment or being asked questions orally. The fact that the learners are being 'watched' and heard by the instructors may be the reason for the high number of learners who consider the virtual classroom reliable.

Regarding the last 2 questions, the data seem to show a discrepancy since more than a third of the respondents consider each of these points only slightly reliable while another third consider them reliable. Several readings could be attributed to this point but, as it would be speculative, we would rather not comment on them.

Analysing the 'reliable' replies, they are relatively consistent for all four questions, between a third and almost half of the learners have given this rating to all the variables. While the 'very reliable' answers are not so clear-cut, when we consider the two ('reliable' and 'very reliable') together, each variable represents over 50%. Thus, we can conclude that, overall, for the learners of UFPUV the control mechanisms used, or that are likely to be used, ensure the feasibility of the assessment process.

5 CONCLUSIONS

The increasing interest of all sectors of society in technology has resulted in the need for resources that will ensure better and more reliable assessment of e-learning courses so they will become socially accepted and those who graduate from online courses may start on the same footing as learners who did F2F courses.

The results of our study demonstrated that the virtual classroom was, for both instructors and learners, the control mechanism that shows the most reliability whether the virtual classroom acts as a 'real' classroom in which both parties are 'face-to-face' via audiovisual means or only the audio means of the virtual classroom is used to question the learner orally.

Although not within the scope of this paper, the interest in e-portfolios as a reliable means of assessment may result in stake-holders and employers seeing e-learning in a different way since e-portfolios provide evidence of students' learning performance and progress, and self-reflection and

self-assessment on their own work. If these portfolios are followed by an oral presentation, instructors will be assessing not only the written work but also the learner's communicative competence.

REFERENCES

[1] D. Fernandes, "Avaliação em Educação: uma discussão de algumas questões críticas e desafios a enfrentar nos próximos anos", *Ensaio: Avaliação e Políticas Públicas em Educação*, vol. 21, no. 78, pp. 11-34, 2013.

[2] D. Boud, "Sustainable assessment: rethinking assessment for the learning society". *Studies in Continuing Education*, vol. 22, no. 2, pp. 151-167, 2000.

[3] K. Scalise & B. Gifford, "Computer-based assessment in E-learning: A framework for constructing "Intermediate Constraint" questions and tasks for technology platforms", *Journal of Technology, Learning, and Assessment*, vol. 4, no. 6, 2006. Retrieved from https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1653/1495

[4] C.P.M. Vleuten, "The assessment of professional competence: Developments, research and practical implications", *Advances in Health Sciences Education*, vol.1, no.1, pp. 41-67, 1996.

[5] P. Race, *The Lecturer's Toolkit:* 4th Edition. London: Routledge, 2015.

[6] H. Abboudi, *et alii*. (2013). "Current status of validation for robotic surgery simulators – a systematic review". *BJU Int*, vol. 111, pp. 194–205, 2013. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1464-410X.2012.11270.x/full

[7] J. Norcini, et al., "Criteria for good assessment: consensus assessment and recommendations from the Ottawa 2010 Conference", *Medical Teacher*, vol. 33, no 3, pp. 206–214, 2011.

[8] P. Race, Making learning happen: A guide for post-compulsory education, 3rd Edition. London: Sage, 2014.

[9] M. Palloff & K. Pratt, Assessing the Online Learner: Resources and Strategies for Faculty. San Francisco: Jossey-Bass, 2009.

[10] L. Eyal, "Digital Assessment Literacy—the Core Role of the Teacher in a Digital Environment", *Educational Technology &Society*, vol.15, no. 2, pp. 37–49, 2012.

[11] T. Palm, "Performance Assessment and Authentic Assessment: A Conceptual Analysis of the Literature", *Practical Assessment Research & Evaluation*, vol. 13, no. 4, 2008. Retrieved from http://pareonline.net/getvn.asp?v=13&n=4

[12] B. Watwood, J. Nugent & W. Deihl, Online Teaching and Learning White Paper: Building from Content to Community: [Re]Thinking the Transition to Online Teaching and Learning. Richmond, Virginia: Virginia Commonwealth University, Center for Teaching Excellence, 2009.

[13] B. Rienties, et al., "Why some teachers easily learn to use a new virtual learning environment: a technology acceptance perspective", *Interactive Learning Environments*, 24(3), pp. 539–552, 2016. Retrieved from http://oro.open.ac.uk/

[14] C. Dalsgaard, "Social Software: E-learning beyond learning management systems", *European Journal of Open, Distance and E-learning (EURODL), 12 July 2006.* Retrieved from http://www.eurodl.org/?p=archives&sp=full&article=228

[15] L. Guardia, G. Crisp & I. Alsina, "Trends and challenges of e-assessment to enhance student learning in higher education". In *Innovative practices for higher education assessment and measurement* (E. Cano & G. Ion eds.), pp. 36–56, Hersey, PA: IGI Global, 2016.

[16] F. Silva, Assessing online learning. *EDULEARN14 Proceedings*, pp. 2432-2440, 2014.

[17] UFPUV - Learning, 2012. Retrieved from http://www.ufpuv.com/metodologia/aprendizagem/?lang=en

[18] UFPUV - Assessment, 2012. Retrieved from http://www.ufpuv.com/metodologia/avaliacao/?lang=en

[19] P. Reis. "e-Learning in the Portuguese Educational Context" in *Trends and Issues in Distance Education: International Perspectives* (L. Visser, Y. Visser, R. Amirault & M. Simonson eds.), pp. 230 - 242. Charlotte, North Carolina: Information Age Publishing, 2012.